Neural Network Based Decision Fusion for Abnormality Detection via Molecular Communications

1st Sinem Nimet Solak
School of Engineering and Informatics
University of Sussex
Brighton, United Kingdom
sinem.solak@sussex.ac.uk

2nd Menguc Oner
School of Engineering and Informatics
University of Sussex
Brighton, United Kingdom
m.m.oner@sussex.ac.uk

Abstract—Abnormality detection is one of the most highly anticipated application areas of Molecular Communication (MC) based nanonetworks. This task entails sensing, detection, and reporting of abnormal changes in a fluid medium that may characterize a disease or disorder using a network of collaborating nanoscale sensors. Existing strategies for such distributed collaborative detection problems require a complete statistical characterization of the underlying communication channel between the sensors and the fusion centre (FC), with the assumption of perfectly-known or accurately estimated channel parameters. This assumption is usually impractical both due to mathematical intractability of the analytical channel models for MC except in a few ideal cases, and the slow and dispersive signal propagation characteristics that make the channel estimation a difficult task even in these ideal cases. This work, for the first time in the literature, proposes to employ a machine learning approach to this task and shows that this approach provides the robustness and flexibility required for practical implementation. We focus on detection based on deep learning, specifically on a feed-forward neural network and a recurrent neural network structure that learn the underlying model from data. This study shows that the proposed decision fusion strategy can perform well without any knowledge of the communication channel.

Index Terms—deep learning, molecular communication, distributed detection, sensor networks

I. Introduction

Nanotechnology has been an attractive and promising research area in the recent years, drawing a considerable research effort from diverse fields of science and engineering. In this regard, nanonetworks, have become an active research area. Conventional communications based on electromagnetic (EM) radiation may not be suitable for nanonetwork applications in aqueous media due to the adverse propagation characteristics of EM waves in water [1]. Molecular Communication (MC) presents a promising alternative for this task by mimicking the naturally evolved communication mechanisms between biological entities at this physical scale. Diffusive MC employs dedicated molecules as information carriers, while relying on the diffusion of these molecules for signal propagation, requiring no infrastructure or additional external energy [2].

This work focuses on the task of abnormality detection, i.e. the detection and reporting of abnormal changes in a fluid environment that may characterize the presence of a disorder using a diffusive MC based nanoscale sensor network. Such distributed detection (DD) problems are encountered in many potential application areas of nanonetworks, e.g. health monitoring, disease diagnosis, targeted drug delivery, environmental sensing and monitoring, contaminant and toxic agent detection, etc. Abnormalities to be detected may be diverse depending on the particular application. Abnormal changes in the concentration of an indicator molecule in the medium, or in the properties of the medium itself, such as the pH value, temperature, pressure, viscosity, etc. are some of the abnormalities that may be of interest in the context of this task [3], [4].

There have been few studies considering abnormality detection with a centralized MC based sensor network architecture based on the assumption that the underlying channel model and its parameters are well known at the receiver side. While [5] considers an AWGN channel model, [6], [7] both propose decision fusion (DF) strategies based on a Poisson model. While allowing the derivation of likelihood ratio based tests (LLR) for DF at the fusion centre (FC), these assumptions fail to take into account the complex and highly dynamic nature of the propagation environments found within living organisms and ecosystems [1]. The signal propagation characteristics in a diffusive MC channel are highly random in nature and depend heavily on a multitude of factors such as the geometry (i.e. the physical bounds) of the fluid medium of propagation, its chemical properties, the chemical characteristics and the physical extent of the transmitters and the receivers, the geometry of the network (i.e. the relative positions of the transmitters and the receivers), and the environmental conditions, such as flow, temperature, viscosity, physical obstacles, etc. (see [8] and the references therein). However, the analytical models for diffusive MC channels proposed in the existing literature are only valid under a set of specific simplifying assumptions, such as an unbounded medium, point transmitters, perfectly spherical absorbing receivers, within a static environment, where the physical and chemical characteristics of the medium remain unchanged during transmission. In practical scenarios

where one ore more of these assumptions don't hold, the channel model may become too complicated, or even mathematically intractable. Even if the channel model is assumed to be known, the estimation of instantaneous channel state information (CSI), i.e. the channel parameters required for detection, is an extremely challenging task, especially due to the extremely low signal propagation speeds and the highly dispersive nature of the MC channel that lead to very long symbol transmission intervals [9]. Hence, the use of methods that require a complete statistical characterization of the signal propagation characteristics in diffusive MC channels may be impractical in application scenarios that are expected to be encountered in MC based nanoscale sensor network implementations.

An emerging alternative approach for communication problems, where an analytical channel model is not mathematically tractable, or too complex to be of practical use, and for cases where a channel model exists but an accurate estimation of its parameters is impractical, is to apply machine learning (ML) techniques that learn from data. Recent works, such as [9] have proposed the use of an end to end machine learning approach for data demodulation in communication problems without any knowledge of the underlying channel model and demonstrate the use of recurrent neural networks (RNN) in the demodulation of sequences of data symbols for point-to point optical- and molecular communications.

In this paper, for the first time in the literature, we propose a deep learning-based end-to end approach to DF in a diffusive MC based nanoscale sensor network. We show that the proposed approach can be used to design DF algorithms performing well under different channel conditions and/or without any instantaneous knowledge on the channel parameters and the statistical characteristics of the sensing model of the participating sensor nodes, providing the robustness and flexibility required for practical implementation of such nanoscale sensor networks. In particular, we investigate the use of feed-forward and recurrent neural networks (FF-NN and RNN respectively) structures through supervised learning to perform the DF. We employ the algorithm proposed in [6] which is based on an approximation of the Log-Likelihood-Ratio (LLR) of the sensor outputs received at the FC, as a benchmark for performance evaluation of the proposed algorithms and show that the proposed methodology leads to a higher and more robust detection performance.

II. System Model and Problem statement

A centralized sensor network with M identical nanoscale sensors transmitting their soft decisions via diffusive MC to a spherical FC in an unbounded 3D environment is considered in this paper as depicted in (Fig 1). Let the hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$ represent the absence and the presence of the abnormality of interest, respectively. The task of the FC is to perform the DF, i.e. to decide for $\mathcal{H}_0$ or $\mathcal{H}_1$ by observing the sensor outputs received via the MC channel. To provide a fair comparison between the proposed machine learning based methodology and the existing model based approaches, we employ the
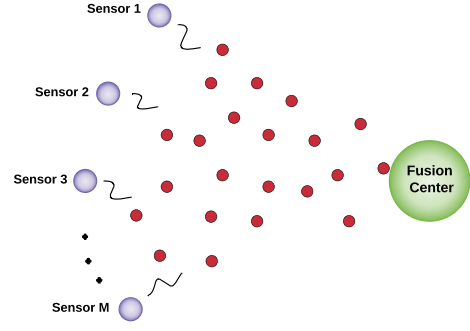

Fig. 1. Sensor network consisting of M nanosensors and FC.

same general and abstract sensing model at the participating sensor nodes, and the same communication model between the sensors and the FC as the method proposed in [6], which is used in this work as a benchmark for performance evaluation. Each of the M sensors is assumed to measure one or more sensing variables, i.e. inputs, and generates a quantized soft output between 0 and 1 representing the sensing decision, $X_m \in \{0, 1/(L-1), 2/(L-1), ...1\}$ where L is the number of quantization levels. The imperfection of the sensors is accounted for by modelling X_m as random variables:

$$X_m \sim \begin{cases} g_0\left(x_m\right), & \text{for } \mathcal{H}_0 \\ g_1\left(x_m\right), & \text{for } \mathcal{H}_1 \end{cases} \quad (1)$$

where $g_i(.)$ denotes the conditional probability mass function of X_m under the hypothesis $\mathcal{H}_i$, $i = 0, 1$. As in [6], we consider a time-slotted transmission with the time slots of duration T, and a total reporting time of $[0, NT]$ where N is the number of time slots. The m'th nanosensor releases $x_m A_{max}$ information carrying molecules at the beginning of each time slot where A_{max} is the maximum number of molecules that a sensor can release and x_m is a realization X_m. These molecules propagate through the fluid medium via diffusion. The FC, a sphere of radius r_r with fully absorbing boundaries, absorbs and counts the information carrying molecules hitting its surface. Channels between the nanosensors and FC are assumed to be statistically independent and identical and the distances between all nanosensors and FC are identical as well. Without loss of generality, we use the Poisson channel model for training and evaluation of our algorithms, which is a commonly employed model in diffusive MC [10]. This approach enables a fair comparison between the proposed methods and the LLR algorithm in [6] employing this channel model, which we use as a benchmark. Let h_k denote the hitting probability. i.e., the probability that a molecule released in the current time slot hits the FC during the k-th next time slot. For a transmitter located at distance r_t from the surface of the FC, h_k is given by [11]:

$$h_k = \begin{cases} \frac{r_r}{r_t + r_r} \operatorname{erfc}\left(\frac{r_t}{\sqrt{4DT}}\right), & \text{if } k = 0 \\ \frac{r_r}{r_t + r_r} \left(\operatorname{erfc}\left(\frac{r_t}{\sqrt{4D(k+1)T}}\right) - \operatorname{erfc}\left(\frac{r_t}{\sqrt{4DkT}}\right)\right), \\ \qquad\qquad \text{if } k \geq 0 \end{cases}$$
$$(2)$$

where D is the diffusion coefficient of the transmitted molecule in the medium and *erfc(.)* is the complementary error function. Let Y_n be an RV representing the number of molecules observed at the FC in the n-th time slot and y_n a realization of RV Y_n. Under the Poisson model, it can be shown that, for a given realization of $X_m = x_m$, $m = 1, \ldots, M$, Y_n is Poisson distributed:

$$Y_n \sim \text{Poisson}\left(J + \sum_{m=1}^{M} \sum_{k=0}^{N} h_k x_m A_{\max} \right), \qquad (3)$$

where J represents the expected number of environmental noise molecules at the FC, and $P(X_m = x_m | \mathcal{H}_i) = g_i(x_m)$, $i = 0, 1$. [6] derives the likelihood function of Y_n over the whole observation period of N samples by averaging (3) over the sensor characteristics,

$$P\left(Y_1 = y_1, \ldots, Y_N = y_n | \text{H}_i \right)$$
$$= \sum_{x_m \in \mathcal{X}} \prod_{n=1}^{N} P\left(Y_n = y_n | \text{H}_i \right) G_i(x_m), \qquad (4)$$

where $G_i(x_m)$ represents the aggregate effect of the individual sensor characteristics given in (1). The LLR test in [6] performs the DF at the FC by using an approximation of (3) employing some limiting assumptions. Clearly, the derivation of the likelihood function requires existence of a parametric model, the presence of a-priori knowledge of a set of channel parameters and of the sensor distributions. Note that even if the channel model is assumed to provide an accurate statistical representation of the signal propagation, these parameters may change over time with environmental conditions, by drifting away of the sensors from the FC, etc. This requires an accurate estimation of these parameters by the FC, which may be infeasible in practice. By considering these parameter variations in the training phase, the proposed NN based approach provides an inherent robustness to such variations.

III. TRAINING THE NEURAL NETWORK DETECTORS

We propose to perform the DF at the FC by using a NN based ML approach through supervised learning. As the first phase of the approach, the detector is trained by using a training dataset. Once the training phase is completed offline, it is no longer a part of the detection procedure.

1 A Feed Forward NN is employed for DF with one input layer consisting of 16 nodes, one fully connected hidden layer consisting of 16 nodes, and one Dense layer as an output layer .
2 A RNN with one Long short-term memory (LSTM) layer consisting of 16 nodes and one Dense layer as an output layer.

We choose the RNN approach which is often used for the tasks that involve sequential inputs, which makes it especially suitable for this DF task. One of the special kind of RNN, LSTM architecture is selected in this study [12]. Let $\mathbf{y}^{(l)} = $

$[y_1^{(l)}, y_2^{(l)}, ..., y_N^{(l)}]$ represent a known signal sequence of length N observed at the FC for the corresponding known hypothesis $\mathcal{H}_i^{(l)}$. The training data consists of L samples of $(\mathbf{y}^{(l)}, \mathcal{H}_i^l)$ pairs.

$$\{(\mathbf{y}^{(1)}, \mathcal{H}_i^{(1)}), (\mathbf{y}^{(2)}, \mathcal{H}_i^{(2)}), \ldots, (\mathbf{y}^{(L)}, \mathcal{H}_i^{(L)})\} \qquad (5)$$

The training data can be generated by various means, e.g. from in vivo measurements, from in vitro experimental data, from particle-based diffusion/reaction simulations of the sensor network, or by generating the corresponding signal sequences using the existing statistical system-and channel models, such as the one described in section II. This makes the proposed approach applicable even if no tractable analytical description is available for the statistical characteristics of the channel. For the sake of a fair comparison between existing model based approaches based on the Poisson channel model, we have chosen to generate the training data set employing the Poisson model from section II, using various values for the parameter set $\zeta_{NN} = [J, r_r, r_t, D]$, (see table I). This makes the resulting DF algorithms robust to possible changes in the system parameters, while allowing DF without requiring parameter estimation, provided that the system parameters remain within the boundaries of the training set. The data set consisting of 100k samples is used to train the proposed FF-NN and RNN detectors that map the received signal $\mathbf{y}^{(l)}$ to a hypothesis $\hat{\mathcal{H}} \in \{\mathcal{H}_0, \mathcal{H}_1\}$. During the training, the known $(\mathbf{y}^{(l)}, \mathcal{H}_i^{(l)})$ pairs are used to find the optimal set of parameters for the corresponding NN structure. Both for the FF-NN and RNN training, we apply back-propagation with binary cross-entropy to minimize the the loss between the actual hypothesis $\mathcal{H}_i$, and estimated hypothesis $\hat{\mathcal{H}}$. Adam optimizer is applied with a learning rate of 10^{-3} to calculate the FF-NN parameters without over-learning. The number of epochs used during training is 100 and the batch size is 10. This optimizer maintains the learning rate for each network weight and makes use of the average of the second moments of the gradients. For the RNN, the RMSoptimizer is employed with a learning rate of 10^{-3} which is a the gradient descent algorithm with momentum which helps to calculate the RNN parameters without over-learning. For the RNN, the same number of training epochs and the same batch size has been employed as for the FF-NN case.

IV. PERFORMANCE EVALUATION

In the following, we evaluate the performance of the proposed NN based detectors via simulations and use the existing LLR detector from [7] as a benchmark. We employ the detection probability $P_d = P(\hat{\mathcal{H}} = \mathcal{H}_1 | \mathcal{H}_1)$ and the false alarm rate $P_f = P(\hat{\mathcal{H}} = \mathcal{H}_1 | \mathcal{H}_0)$ as our main performance criteria. System parameters used in the simulation setup are presented in (Table I). In the training and test data, we use $N=4$, and J varies depending on the SNR level, which we define as $SNR = \frac{A_{max}}{J}$. The conditional pmf's of the sensors

used both on the simulations and the training are given as:

$$g_0(x_m) = \frac{\exp\left(-6.5 x_m\right)}{\sum_{x_m \in \mathcal{X}} \exp\left(-6.5 x_m\right)}, \tag{6}$$

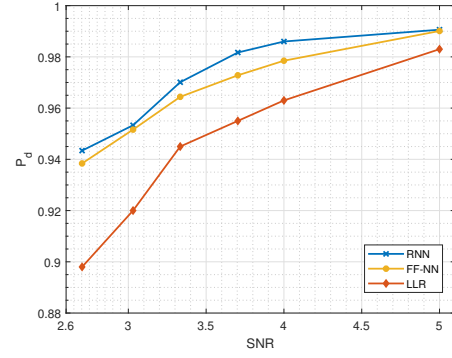$$g_1(x_m) = \frac{\exp\left(7.5 x_m\right)}{\sum_{x_m \in \mathcal{X}} \exp\left(7.5 x_m\right)}. \tag{7}$$

for $x_m \in \mathcal{X} = \{0, 1/(L-1), 2/(L-1), ..., 1\}$ with $L = 4$. Figs 2 (a) and (b) exhibit the detection performance of the proposed RNN and FF-NN detectors, and of the existing LLR detector from [6] for a fixed false alarm rate $P_f = 0.05$ and 0.075 respectively. Clearly, both of the proposed NN based detectors outperform the LLR detector, which is due to the fact that the LLR employs only an approximation of the likelihood function, whereas the proposed detectors both learn from actual data resulting in a better approximation to the impractical optimal detector. It can be seen from Fig 2 that, between both of the proposed detectors, the RNN approach outperforms the FF-NN. Therefore, in the rest of the paper we focus on the evaluation of the RNN detector. This is due to the well known capability of this type of neural networks for taking into account the temporal structure of the data, which makes RNNs especially suitable for this DF task. Fig 3 compares the receiver operating characteristics (ROC) of the RNN and the LLR detectors for two different SNR levels. As in the previous case, the RNN detector performs better than LLR detector in each case. Note that that the proposed detector is robust enough to changes in noise level, since the training data set includes data with different SNR levels.

Next, we study the effect of the sensor to FC distance r_t, which is a crucial parameter for the channel characteristics, on the detection performance of the proposed RNN detector, and compare it to that of the benchmark LLR. We consider 3 different values for the P_f and $SNR = 2$. Fig 4 clearly shows that the proposed RNN detector for DF outperforms the LLR, without requiring any prior knowledge, or estimation, of the parameter r_t. A similar effect is observed in Fig 5, where we investigate the effect of another key factor in signal propagation, which is the diffusion coefficient, D. The performance of RNN detector is presented for $D=50$ $\mu m^2/s$ and $D=79.4$ $\mu m^2/s$ representing the diffusion coefficient of ionic calcium in cytoplasm and the diffusion coefficient of human insulin hormone-like molecules in a blood-like fluid, respectively.
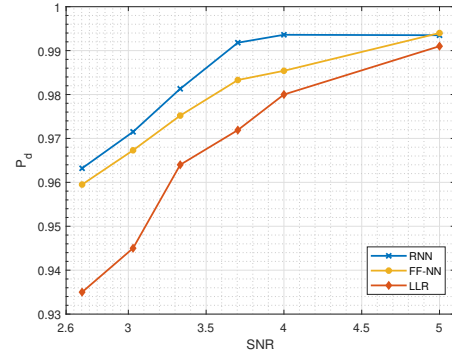

(a) Probability of detection vs SNR (P_f=0.05).

(b) Probability of detection vs SNR (P_f=0.075).

Fig. 2. Probability of detection vs Signal to noise ratio (SNR) for P_f=0.05 (top) and P_f=0.075 (bottom).
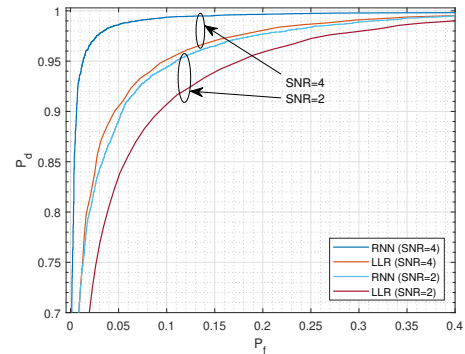

Fig. 3. ROC curve of RNN and LLR detector for different SNR levels.

TABLE I
LIST OF PARAMETERS USED IN THE TRAINING AND THE SIMULATIONS

Parameter	Conf. 1
A_{max}	100
Time slot duration	$T = 70ms$
Number or sensors	4
Number of quantization levels	$L = 4$
Receiver radius	$4\mu m$
Distance btw sensors and receiver	r_t=4,5,6,7,8,9μm
SNR levels	2-5
Diffusion coefficient	D=50, 79.4 $\mu m^2/s$

The simulation results indicate that the RNN detector performs well and outperforms LLR detector for different CSI parameters, SNR, r_t, D. Note that in MC these parameters may change rapidly due to the nature of the environment. Therefore, robustness is a key feature in MC applications besides good detection performance. Clearly, training the RNN detector using a data set containing data produced under different channel conditions provides a robustness to the detector that is a key requirement in actual practical applications.
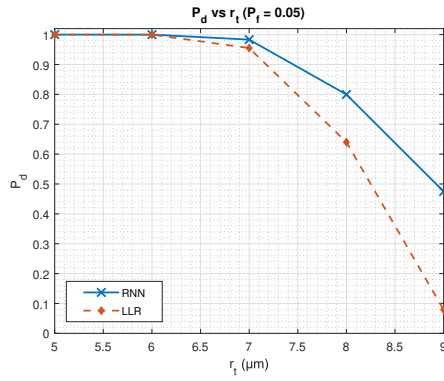
(a) Probability of detection vs r_t (P_f=0.05).
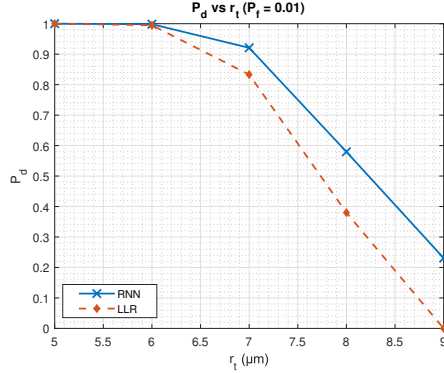
(b) Probability of detection vs r_t (P_f=0.01).
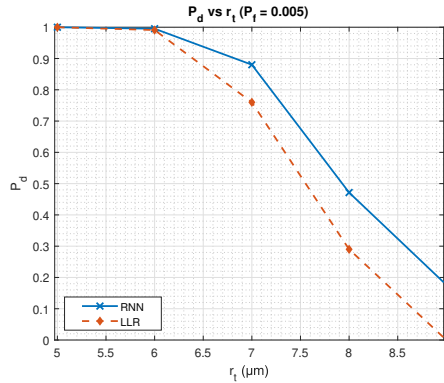
(c) Probability of detection vs r_t (P_f=0.005).

Fig. 4. The P_d performance of the RNN detector and LLR detector for different P_f considering different r_t values (SNR=2).
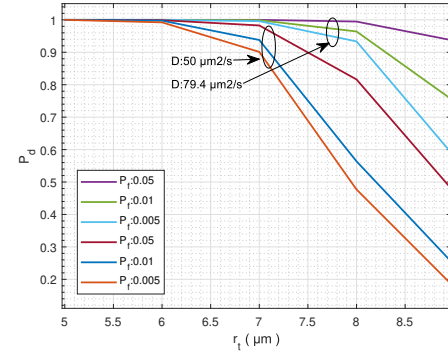

Fig. 5. RNN performance for different Diffusion coefficient (D) values (SNR=2).

particle based diffusion simulators, etc, without requiring an analytical statistical description of the channel model. Thus, this approach leads to DF methods which don't require any analytical channel model, instantaneous channel state information or knowledge of other system parameters such as the statistical characteristic of the sensors, even the number of sensors, etc. The presented results indicate that the novel ML based DF approach proposed in this work provides good detection performance along with a robustness and flexibility that is required for practical implementations of such MC based nanoscale sensor networks.

V. CONCLUSION

In this paper, we have demonstrated the use of machine learning techniques in the design of DF algorithms for a centralized nanoscale sensor network for detecting abnormalities in a fluid environment using MC. We have investigated the use of two different NN structures for detection, the FF-NN and the RNN. As expected, the RNN based approach has proven to be more suitable for the proposed task due to the sequential nature of the data encountered in MC and its temporal structure, which the RNN is able to make use of. In our proposed approach, ANN detectors can be trained directly with data obtained from mathematical models, or, alternatively, from in vivo measurements, in vitro experiments,

REFERENCES

[1] T. Nakano, A. W. Eckford, and T. Haraguchi, *Molecular Communication*. Cambridge University Press, 2013.
[2] I. F. Akyildiz, F. Brunetti, and C. Blázquez, "Nanonetworks: A new communication paradigm," *Computer Networks*, 2008.
[3] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Reference Format*, vol. 41, no. 15, 2009. [Online]. Available: http://doi.acm.org/10.1145/1541880.1541882
[4] S. Ghavami, F. Lahouti, and A. Masoudi-Nejad, "Modeling and analysis of abnormality detection in biomolecular nano-networks," *Nano Communication Networks*, vol. 3, no. 4, pp. 229–241, 2012. [Online]. Available: http://dx.doi.org/10.1016/j.nancom.2012.09.008
[5] S. Ghavami and F. Lahouti, "Abnormality Detection in Correlated Gaussian Molecular Nano-Networks: Design and Analysis," *IEEE Transactions on Nanobioscience*, 2017.
[6] R. Mosayebi, V. Jamali, N. Ghoroghchian, R. Schober, M. Nasiri-Kenari, and M. Mehrabi, "Cooperative Abnormality Detection via Diffusive Molecular Communications," *IEEE Transactions on Nanobioscience*, 2017.
[7] R. Mosayebi, W. Wicke, V. Jamali, A. Ahmadzadeh, R. Schober, and M. Nasiri-Kenari, "Advanced Target Detection via Molecular Communication," *2018 IEEE Global Communications Conference, GLOBECOM 2018 - Proceedings*, pp. 1–7, 2018.
[8] N. Farsad, H. B. Yilmaz, A. Eckford, C. B. Chae, and W. Guo, "A comprehensive survey of recent advancements in molecular communication," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 3, pp. 1887–1919, 2016.
[9] N. Farsad and A. Goldsmith, "Neural network detection of data sequences in communication systems," *IEEE Transactions on Signal Processing*, vol. 66, no. 21, pp. 5663–5678, 2018.
[10] C.-B. Chae and H. B. Yilmaz, "Arrival modelling for molecular communication via diffusion," *Electronics Letters*, 2014.
[11] H. B. Yilmaz, A. C. Heren, T. Tugcu, and C. B. Chae, "Three-dimensional channel characteristics for molecular communications with an absorbing receiver," *IEEE Communications Letters*, 2014.
[12] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.